

Question Paper

Exam Date & Time: 29-May-2023 (10:00 AM - 01:00 PM)



MANIPAL ACADEMY OF HIGHER EDUCATION

Manipal School of Information Sciences (MSIS), Manipal
Second Semester Master of Engineering - ME (Cloud Computing) Degree Examination - May 2023

Data Warehousing and Data Mining Elective -2 [CDC 5232]

Marks: 100

Duration: 180 mins.

Monday, May 29, 2023

Answer all the questions.

- 1) Explain Conceptual Structure of Data Warehouse and discuss Classification of Data Warehouses? (10)
- 2) Explain why Data Preprocessing is required and summarize the major tasks in the Data Preprocessing? (10)
- 3) Summarize Attribute Subset Selection and discuss a) Feature selection b) List Heuristic methods c) discuss Decision tree induction? (10)
- 4) Illustrate and Generate the frequent item sets using the Apriori Algorithm for the following database with (10)

Tid	Items
10	A, C, D
20	B, C, E
30	A, B, C, E
40	B, E

$$\text{Sup}_{\min} = 2$$

- 5) Examine and Describe briefly Two ways of Rule Constraints to be mined? Summarize five types of constraints? (10)
- 6) Differentiate and describe what is Decision Tree Induction? Generate the Decision tree for the following dataset "Training data set for buys_computer"? (10)

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

- 7) Analyze and Explain SVM Algorithm in terms of Hyper-Plane and Support-Vectors and discuss any two kernels used in SVM algorithm? (10)
- 8) Identify and discuss How to Measure the Quality of Clustering and List and briefly write on Considerations/Criteria's for Cluster Analysis? (10)
- 9) Cluster the following 8 points (with (X, Y) representing locations) into three Clusters A1(2,10), A2(2,5), A3(8,4), A4(5,8), A5(7,5), A6(6,4), A7(1,2), A8(4,9). Initial cluster centers are: A2, A4 & A8 ? (10)
- 10) Name and Discuss at least 5 Advantages of cloud computing? (10)

-----End-----